

Томпсона відповідає побудові інтегрального показника за методикою у роботі [4]. Отже, за цією моделлю у межах незначних коливань можна моделювати вплив основних важелів освітнього ринку на показники ефективності роботи ВНЗ на освітньому ринку праці України.

Література

1. Ранжування університетів – крок до відкритості та прозорості вищої освіти // Ранжування вищих навчальних закладів III–IV рівнів акредитації // Статистика МОН відповідно до розпорядження Кабінету Міністрів України № 208-р від 1.04.2004 р. «Про схвалення Концепції створення системи рейтингової оцінки регіонів, галузей національної економіки, суб'єктів господарювання».
2. Жуковська О. А. Інтервальні моделі прийняття колективних рішень в умовах ризику // Автореф. дис. к. ф.-м. н., Ін-т матем. НАН України. – К, 2006. – 19 с.
3. Кендэл М. Ранговые корреляции. ЗСИ; пер. с англ.; под ред. Е. М. Четыркина, Р. М. Энтова. – М.: Статистика, 1975. – 212 с.
4. Матусов Ю. П., Дмитрієв Д. С. Оцінки індикаторів інтегрального показника у дослідженні ринку праці України // Актуальні проблеми економіки та управління: зб. наук. пр. – К.: НТУУ «КПІ», 2008. – Вип. 2. – 188 с. – С. 159–164.
5. Багриновский К. А. Имитационные системы принятия экономических решений / К. А. Багриновский, Т. И. Конник, М. Р. Левинсон и др. – М.: Наука, 1989. – 162 с.: – ил. – 15000 экз. – ISBN 5-02-011880-X4.
6. Айвазян С. А., Мхитарян В. С. Прикладная статистика и основы эконометрики / С. А. Айвазян, В. С. Мхитарян. – М.: ЮНИТИ, 1998. – 1022 с. – 7000 экз. – ISBN 5-238-00305-6.

УДК 331.211:519.86

Т. А. Дунаєва,
к.ф.-м.н., доц.
Т. О. Проноза,

Національний технічний університет України «КПІ»

ЗАСТОСУВАННЯ НЕЙРОННИХ МЕРЕЖ У СИСТЕМІ ЕЛЕКТРОННИХ ПЛАТЕЖІВ

Для виявлення можливих шахрайських операцій із платіжними картками запропоновано застосовувати метод карт Кохонена, що самоорганізуються, який ґрунтується на теорії нейронних мереж і належить до алгоритмів навчання без учителя. Цей метод розглянуто у двох випадках, коли нейрони вихідного шару розташовуються у вузлах двовимірної сітки із прямокутними і шестикутними осередками.

In this article for the exposure of possible knavish operations with pay cards we suggest to apply the method of maps of Kohonen that self organization, which is based on the theory of neuron networks and behaves to the algorithms of studies without a teacher. We examine this method in two cases, when: the neurons of initial layer are disposed in the knots of two-dimensional net with rectangular and hexagonal cells.

Ключові слова: транзакція, нейронна мережа, карта Кохонена, профайл, моніторинг, процесинговий центр, прямокутні осередки, шестикутні осередки.

Вступ. Нині вже немає сумнівів у тому, що ринок пластикових платіжних засобів є одним із найбільш динамічно зростаючих сегментів української економіки. Буквально за один рік електронні платежі увійшли в наше життя і зайняли там досить міцні позиції.

Незважаючи на активний розвиток платіжних систем, український ринок ще далеко не наповнений. Проте темпи росту дають можливість зробити висновок, що незабаром розрахунки за допомогою банківських платіжних карток стануть дійсно масовим явищем. Але разом з тим бізнес у сфері грошового обігу з використанням банківських ПК постійно є мішенню для злочинних зазіхань [1].

Платіжні системи і банки для ефективного виявлення й запобігання шахрайствам із пластиковими картками використовують автоматизовані системи моніторингу, які в загальному потоці транзакцій виявляють підозрілі на шахрайство, попереджають про це аналітиків і надають їм інформацію, необхідну для проведення розслідування й ухвалення рішення [2]. Такі системи дають можливість банкам емітентам і еквайерам зміцнювати захист від шахрайства, знижувати ризики й втрати як самих банків, так і їх клієнтів [3]. Побудова таких автоматизованих систем для процесингового центру (ПЦ) і банків вимагає розробки спеціальних моделей, методів і правил аналізу. У цій роботі ми розглядаємо метод карт Кохонена, що самоорганізуються, який ґрунтується на теорії нейронних мереж і відноситься до алгоритмів навчання без учителя для виявлення можливих шахрайських операцій із платіжними картками. Розглядаємо його у двох аспектах: нейрони вихідного шару розташовуються у вузлах двовимірної сітки із прямокутними і шестикутними осередками.

Постановка завдання. Нехай $C_n = \{c_1, \dots, c_k, \dots, c_{k_n}\}$ – множина записів у базі даних про платіжні картки, де $c_k = (c_1^k, c_2^k, \dots, c_s^k)$ – запис у базі даних з відомостями про карту c_k ; c_1^k – унікальний номер платіжної картки.

Позначимо множину транзакцій, виконаних у платіжній системі до деякого моменту часу t_n через $X_n = \{x^1, \dots, x^i, \dots, x^n\}$, а $x^{n+1}, x^{n+2}, \dots, x^{n+k}$ – нові транзакції, зроблені після моменту часу t_n й до t_{n+k} .

Завдання виявлення шахрайських транзакцій у платіжній системі полягає в тому, щоб при надходженні у процесинговий центр кожної нової транзакції $x^{n+1} = (x_1^{n+1}, \dots, x_j^{n+1}, \dots, x_m^{n+1})$, на основі інформації, яка міститься у базі даних C_n про платіжні картки і X_n раніше виконані транзакції, класифікувати транзакцію x^{n+1} , тобто визначити клас, до якого вона належить (законна (*legal*) або шахрайська (*fraud*)).

Методологія. За Кохоненом [4], нейромережа має всього два шари: один вхідний шар з числом нейронів, що дорівнює числу входів, і один схований вихідний шар нейронів, складений з радіальних нейронів упорядкованої структури (вихідний шар називають також шаром топологічної карти).

Нейрони вихідного шару розташовуються у вузлах двовимірної сітки із прямокутними або шестикутними осередками (рис. 1). При цьому нейрони також взаємодіють один з одним. Величина цієї взаємодії визначається відстанню між нейронами на карті.

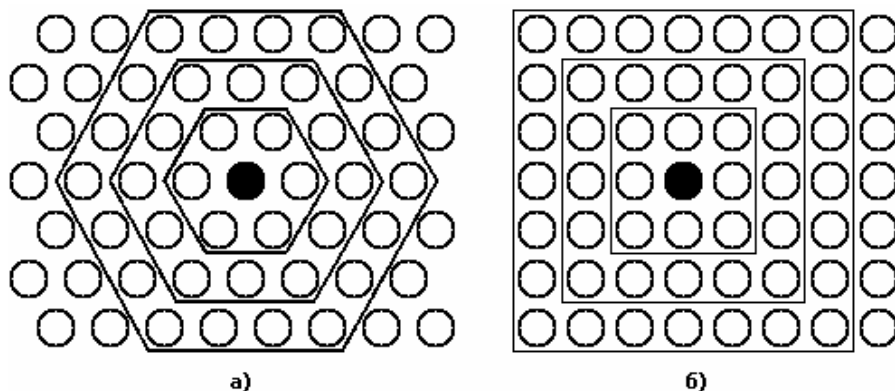


Рис. 1. Відстань між нейронами на карті для шестикутної (а) і чотирикутної (б) сітки

При цьому кількість нейронів у сітці визначає ступінь деталізації результату роботи алгоритму, і в остаточному підсумку від цього залежить точність узагальнюючої здатності карти [3].

У процесі навчання на вхід нейромережі послідовно подаємо навчальні приклади, вибрані випадковим чином з вихідного набору даних. Після подачі чергового прикладу $x = (x_1, x_2, \dots, x_n)$ визначаємо найбільш схожий на нього вектор коефіцієнтів нейронів $w^i = (w_1^i, w_2^i, \dots, w_n^i)$. При цьому обираємо нейрон-переможця, що найбільш схожий на вектор входів. Під схожістю розуміємо відстань між векторами. У цій задачі розглядаємо евклідовий простір, відстань між векторами (x) і (y) визначається як

$$d(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}.$$

Таким чином, нейроном-переможцем (c) є такий нейрон, відстань між вектором ваг й вхідним вектором якого є мінімальною:

$$c: \|x - w^c\| = \min_i \{\|x - w^i\|\}.$$

Після того, як знайшли нейрон-переможця, коректуємо ваги нейромережі. Модифікація вагових коефіцієнтів нейрона-переможця і нейронів з його околиці відбувається в декілька етапів.

1. Знаходимо відстані між нейронами: r_{cj} – відстань між нейронами c та j (c – номер нейрона – переможця). Цю величину визначаємо топологією осередків карти Кохонена; r_{cj} – елемент, який знаходиться в c -му рядку і j -му стовпці матриці R , яка має вигляд

$$R = \begin{pmatrix} d(Y_{00}; Y_{00}) & d(Y_{00}; Y_{01}) & \dots & d(Y_{00}; Y_{0m}) & \dots & d(Y_{00}; Y_{n0}) & d(Y_{00}; Y_{n1}) & \dots & d(Y_{00}; Y_{nm}) \\ d(Y_{01}; Y_{00}) & d(Y_{01}; Y_{01}) & \dots & d(Y_{01}; Y_{0m}) & \dots & d(Y_{01}; Y_{n0}) & d(Y_{01}; Y_{n1}) & \dots & d(Y_{01}; Y_{nm}) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ d(Y_{0m}; Y_{00}) & d(Y_{0m}; Y_{01}) & \dots & d(Y_{0m}; Y_{0m}) & \dots & d(Y_{0m}; Y_{n0}) & d(Y_{0m}; Y_{n1}) & \dots & d(Y_{0m}; Y_{nm}) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ d(Y_{n0}; Y_{00}) & d(Y_{n0}; Y_{01}) & \dots & d(Y_{n0}; Y_{0m}) & \dots & d(Y_{n0}; Y_{n0}) & d(Y_{n0}; Y_{n1}) & \dots & d(Y_{n0}; Y_{nm}) \\ d(Y_{n1}; Y_{00}) & d(Y_{n1}; Y_{01}) & \dots & d(Y_{n1}; Y_{0m}) & \dots & d(Y_{n1}; Y_{n0}) & d(Y_{n1}; Y_{n1}) & \dots & d(Y_{n1}; Y_{nm}) \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ d(Y_{nm}; Y_{00}) & d(Y_{nm}; Y_{01}) & \dots & d(Y_{nm}; Y_{0m}) & \dots & d(Y_{nm}; Y_{n0}) & d(Y_{nm}; Y_{n1}) & \dots & d(Y_{nm}; Y_{nm}) \end{pmatrix} \begin{matrix} N(Y_{01}) \\ N(Y_{00}) \\ N(Y_{0m}) \\ N(Y_{00}) \\ N(Y_{01}) \\ N(Y_{0m}) \end{matrix}$$

$$\begin{matrix} N(Y_{00}) & N(Y_{01}) & N(Y_{0m}) & N(Y_{n0}) & N(Y_{n1}) & N(Y_{nm}) \end{matrix}$$

Якщо топологія карти прямокутна, то $d(Y_{ij}; Y_{k,s})$ обчислюємо за формулою: $d(Y_{ij}; Y_{k,s}) = \sqrt{(i-k) + (j-s)}$.

Якщо топологія карти шестикутна, то $d(Y_{ij}; Y_{k,s})$ обчислюємо за формулою $d(Y_{ij}; Y_{k,s}) = \sqrt{[(i+\alpha)-(k+\beta)]^2 + [(j+\gamma)-(s+\lambda)]^2}$,

де

$$\alpha = \begin{cases} 0, \frac{N(Y_{i,j})+1}{2} \text{ (не ціле число)} \\ 0,5, \frac{N(Y_{i,j})}{2} \text{ (ціле число)} \end{cases}; \beta = \begin{cases} 0, \frac{N(Y_{i,j})+1}{2} \text{ (не ціле число)} \\ 0,5, \frac{N(Y_{k,s})+1}{2} \text{ (ціле число)} \end{cases};$$

$$\gamma = \begin{cases} 1, \frac{N(Y_{i,j})+1}{2} \text{ (не ціле число)} \\ \frac{\sqrt{3}}{2}, \frac{N(Y_{i,j})+1}{2} \text{ (ціле число)} \end{cases}; \lambda = \begin{cases} 1, \frac{N(Y_{k,s})+1}{2} \text{ (не ціле число)} \\ \frac{\sqrt{3}}{2}, \frac{N(Y_{k,s})+1}{2} \text{ (ціле число)} \end{cases}.$$

2. Знаходимо радіус навчання $\delta(t)$:

$$\delta(t) = \delta(0) + (u-1) \Delta\delta,$$

де

$$\Delta\delta = \frac{\delta_{end} - \delta(0)}{q-1},$$

де $\delta(0)$ – початкове значення радіуса навчання (задаємо спочатку навчання як параметр);

δ_{end} – кінцеве значення радіуса навчання (задаємо спочатку навчання як параметр);

q – загальна кількість транзакцій для навчання.

3. Знаходимо функцію сусідства $h(r_{c,j}, t)$:

$$h(r_{c,j}, t) = \exp \left\{ -\frac{r_{c,j}^2}{2\delta^2(t)} \right\}.$$

4. Знаходимо коефіцієнт корегування ваг $h_{cf}(u)$:

$$h_{cf}(t) = h(r_{c,j}, t) \cdot a(t),$$

де $a(u)$ – функція швидкості навчання $a(t) = a_0 \cdot \left(1 - \frac{t}{q}\right)$,

де a_0 – початкове значення швидкості навчання (задаємо спочатку навчання як параметр);

a_{end} – кінцеве значення швидкості навчання (задаємо спочатку навчання як параметр).

5. Формуємо нову матрицю вагових коефіцієнтів:

$$w^i(t+1) = w^i(t) + h_{cf}(t) \cdot [x(t) - w(t)],$$

де t – номер епохи (крок навчання);

$x(t)$ – випадковим чином обраний на кроці t вектор з навчальної вибірки;

$w^i(t)$ – вектор ваг i -го нейрона на кроці t ;

$h_{ci}(t)$ – функція сусідства нейронів, що являє собою незростаючу функцію від часу й відстані між нейроном переможцем і сусідніми нейронами в сітці.

Результатом навчання є ваги входних з'єднань нейронів, які щонайкраще групують підмножини входних даних, утворюючи їх прототипи-профайли:

$$W = \begin{pmatrix} w_1^1 & w_1^2 & \dots & w_1^d \\ w_2^1 & w_2^2 & \dots & w_2^d \\ \dots & \dots & \dots & \dots \end{pmatrix}.$$

Після проведення процесу навчання кожену нову транзакцію x^{n+1} , що здійснюється за картою c_k , перевіряємо на відповідність профайлу W_{c_k} .

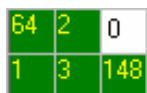
Ступінь подібності $\delta(x^{n+1}, W_{c_k})$ транзакції $x^{n+1} = (x_1^{n+1}, \dots, x_m^{n+1})$ профайлу W_{c_k} будемо визначати як відхилення вектора $p^{n+1} = \phi(x^{n+1})$ від найближчого осередку карти W_{c_k} , тобто як мінімальну з відстаней між вектором $p^{n+1} = (p_1^{n+1}, \dots, p_M^{n+1})$ і векторами ваг нейронів мережі w^1, \dots, w^d :

$$\delta(x^{n+1}, W_{c_k}) = \min_{i=1; 2; \dots; d} \|p^{n+1} - w^i\|.$$

Залежно від специфіки простору даних і особливостей очікуваних результатів використовуються різноманітні (квадрат Евклідової відстані, манхтенська метрика, відстань Чебишева і т. д.) метрики.

Ми скористаємось Евклідовою метрикою:

$$\|p^{n+1} - w^i\| = \sqrt{\sum_{k=1}^M (p_k^{n+1} - w_k^i)^2}.$$



а



б

Рис. 2. Карта Кохонена:
а – прямокутна;
б – шестикутна сітка

до процесингового центру, максимальнодопустиму максимальну похибку, то того, що вона є транзакцією, яка перевищує

Для моделі поведінки Кохонена з прямокутними і зображені кластери, типової поведінки власника

З карти кластерів в обох карти має два яскраво транзакції», а також декілька при побудові карти з відбулась більш детально, рожнім.

Після того, як системи було подано дві нові транзакції.

Першою на вхід до системи було подано типову (законну) транзакцію (рис. 3). Зауважимо, що транзакція є законною, коли різниця між допустимими відхиленнями і відхиленнями нової транзакції дає додатні числа. Чим менші відхилення, тим більша вірогідність того, що транзакція є законною.

Результати дослідження. Застосування запропонованого методу було реалізовано на прикладі за допомогою програмного продукту Delphi 7. Для оцінювання транзакцій було обрано такі показники: x_1 – дата та час транзакції, x_2 – тип транзакції (АТМ, POS), x_3 – тип ретейлера, x_4 – код транзакції, x_5 – валюта, x_6 – сума транзакції, x_7 – країна терміналу, x_8 – місто терміналу.

Для того, щоб перевірити точність роботи цієї моделі і порівняти її роботу у двох аспектах: нейрони вихідного шару розташовуються у вузлах двовимірної сітки із прямокутними і шестикутними осередками, – у вибірці для навчання розглядалося 218 транзакцій, із них 24 тестові транзакції, що характеризували поведінку власника карти (дані були змодельовані).

Для карти було побудовано модель типової поведінки власника карти. У результаті проведених розрахунків було знайдено такі дані: $averrl$, $averrt$ – середньодопустима похибка у навчальній і тестовій множинах відповідно, $maxerrl$, $maxerrt$ – максимальнодопустима похибка у навчальній і тестовій множинах відповідно.

Розрахунок відхилення транзакції				
	avverl	maxerrl	avvert	maxerrt
Карта :	0.9428441	3.7115973	1.8694986	3.4521453
Транз-я :	0.0357792	0.0357792	0.0357792	0.0357792
Відхи-я :	0.9070648	3.675818	1.8337193	3.416366

а

Розрахунок відхилення транзакції				
	avverl	maxerrl	avvert	maxerrt
Карта :	0.8853021	3.0159182	1.0359861	2.869473
Транз-я :	0.0319258	0.0319258	0.0319258	0.0319258
Відхи-я :	0.8533763	2.9839924	1.0040603	2.8375472

б

Рис. 3. Результат перевірки типової транзакції для прямокутної (а) і шестикутної (б) сіток

Нові транзакції, які будуть надходити не повинні перевищувати середню та похибку. Якщо транзакція перевищує вона вважається підозрілою і вірогідність шахрайською, досить велика порівняно з лише середню похибку.

власника карти були побудовані карти шестикутними осередками. На рис. 2 отримані в результаті побудови моделі карти для обох випадків.

випадках помітно, що поведінка власника виражені типи, які можна назвати «Типові «Нетипових транзакцій». Відзначимо, що шестикутними осередками кластеризація оскільки жоден з осередків не лишився по-

закінчився процес навчання, на вхід до

Проаналізуємо отримані результати (рис. 3). Першу транзакцію в обох випадках система визнала законною, що повністю відповідає дійсності.

Другою на вхід до системи було подано нетипову (підозрілу на шахрайську) транзакцію (рис. 4). Зауважимо, що транзакція є підозрілою на шахрайську, коли: середні і максимальні відхилення нової транзакції значно перевищують середні і максимальні допустимі відхилення (виділено червоним кольором). Чим більші відхилення, тим більша вірогідність того, що транзакція є підозрілою на шахрайську.

Проаналізуємо отримані результати. Другу транзакцію в обох випадках система визнала підозрілою на шахрайську, що повністю відповідає дійсності (рис. 4). Це транзакція потребує більш детального аналізу з боку аналітика, який повинен прийняти рішення щодо законності нової транзакції. Якщо транзакція, на думку аналітика, є законною, то її приймає система і вносить корективи у створений профайл власника карти, якщо ж вона визнається шахрайською, то система заблокує картку, за якою здійснювалась шахрайська транзакція.

Порівнюючи значення, отримані у випадку з прямокутною і шестикутною сіткою, можна зробити висновок, що у другому випадку отримані результати є більш точними, оскільки відхилення від усіх чотирьох похибок є більшим, ніж у першому випадку.

Можна зробити висновок, що в обох випадка система працює вірно, але при побудові шестикутної сітки результати є більш точними.

Висновки. У роботі запропоновано новий підхід для виявлення підозрілих на шахрайство транзакцій у платіжній системі на основі нейромережного методу карт Кохонена, що самоорганізуються. Цей метод має такі переваги: його зручно використовувати навіть у тих випадках, коли немає попередньої інформації про спостереження за транзакціями з різних класів; достатньо мати інформацію лише з цього класу транзакцій (законних або шахрайських); метод карт Кохонена дає змогу визначити як вже відомі, так і раніше незафіксовані типи шахрайства.

Розглянувши у роботі два підходи до побудови карти Кохонена, а саме коли нейрони вихідного шару розташовуються у вузлах двовимірної сітки із прямокутними осередками та коли нейрони вихідного шару розташовуються у вузлах двовимірної сітки із шестикутними осередками, ми дійшли висновку, що обидва підходи працюють вірно й адекватно; при побудові шестикутної сітки результати є більш точними, оскільки для неї відстань між нейронами більше збігається з евклідовою відстанню, ніж для чотирикутної сітки.

У результаті на основі побудованої моделі виявлення шахрайства з пластиковими картками стає можливим оцінка транзакцій платіжних карт та прийняття рішень відносно законності даних транзакцій. Адекватність моделі експериментально підтверджено.

Література

1. Вертузаев М. С., Кондратьев Я. Ю., Пугачев С. Е. и др. Способы совершения преступлений с использованием банковских платежных карт // Информационные технологии та захист інформації: зб. наук. праць. – Запоріжжя: Юридичний ін-т МВС України, 1999. – Вип. 3. – № 1. – С. 50–67. – ISBN 966-95343-1-3.
2. Гинзбург А. И. Пластиковые карты [Текст]. – СПб.: Питер, 2004. – 128 с. – 2000 экз. – ISBN 546-90016-5-2.
3. Рубинштейн Т. Б., Мирошкина О. В. Пластиковые карты [Текст]. – СПб.: Питер, 2005. – 416 с. – 1000 экз. – ISBN 5-85438-118-4.
4. Kohonen T., Self-Organizing Maps (Third Extended Edition) [Text]. – New York, 2001. – 501 p. – ISBN 3-540-67921-9.
5. Быстров Л. В., Воронин А. С., Гамольский А. Ю. и др. Пластиковые карты [Текст]. – СПб.: Питер, 2007. – 624 с. – 1000 экз. – ISBN 5-93306-066-6.

Розрахунок відхилення транзакції				
	avvertl	maxvertl	avvert	maxvert
Карта :	0,9428441	3,7115973	1,8694986	3,4521453
Транз-я :	15,7123206	15,7123206	15,7123206	15,7123206
Відхил-я :	14,7694765	12,0007233	13,842822	12,2601753

a

Розрахунок відхилення транзакції				
	avvertl	maxvertl	avvert	maxvert
Карта :	0,8853021	3,0159182	1,0359861	2,869473
Транз-я :	16,120732	16,120732	16,120732	16,120732
Відхил-я :	15,2354299	13,1048138	15,084745	13,251259

b

Рис. 4. Результат перевірки транзакції підозрілої на шахрайство для прямокутної (a) та шестикутної (б) сіток